D1.2 Progress report on the implementation of a secure registry-to-registry communication system

Grant Agreement number: 101015514
Project acronym: SAVDON
Work Package number: WP1

Organisation: **World Marrow Donor Association (WMDA)**
LEAR: Lydia Foeken
Project coordinator: Lydia Foeken
Tel: 0031 88 505 7900
E-mail: lydia.foeken@wmda.info
Organisation website address: www.wmda.info

Co-funded by the Health Programme of the European Union

"This Deliverable D1.2 of an activity received funding under an operating grant

Table of Contents

Abbreviations

API = Application Programming Interface
BMDW = Bone Marrow Donors Worldwide
CBB = Cord Blood Bank
CBU = Cord Blood Unit
CPU = Central Processing Unit
DKMS = Deutsche Knochenmark Spenderdatei
DMZ = Demilitarized zone (sometimes referred to as a perimeter network or screened subnet)
EMDIS = European Marrow Donor Information System
FHIR = Fast Healthcare Interoperability Resources
GDPR = General Data Protection Regulation
HL7 = Health Level 7
HLA = Human Leukocyte Antigen
IaaS = Infrastructure as a service
ICT = Information and Communications Technology
IDM = Infectious Disease Markers
IP = Internet Protocol
JSON = JavaScript Object Notation
NoSQL = Stores information in JSON documents instead of columns and rows used by relational databases
OWASP = Open Web Application Security Project
PaaS = Platform as a service
PHP = Hypertext Pre-processor
POC = Proof of concept
QA = Quality Assurance
RDP = Remote Desktop Protocol
SaaS = Software as a service
SIEM = Security Information and Event Management
WAF = Web Application Firewall
WMDA = World Marrow Donor Association
ZKRD = Zentrales Knochenmarkspender-Register Deutschland

1. Introduction

This progress report on the implementation of a secure registry-to-registry communication system details the process and achievements the WMDA and international partners, including EU member states, have made towards implementing a secure registry-to-registry communication system in 2021. This report allows analysis of progress by all parties involved and will enable them to focus their efforts on areas of identified weaknesses. In addition, this report includes an overview of the complete 3-year project from 2019 to 2021 with a high-level strategy for future development, as well as support and maintenance of the existing solution.

This Deliverable *D1.2 Progress report on the implementation of a secure registry-to-registry communication system* is part of the 2021 work programme of the World Marrow Donor Association (WMDA) for the EU Third Health Programme (2014-2020).

1.1 Background

WMDA took over the operation of the Search & Match Service (known as BMDW) on January 1, 2017. At that time phase 1 had just been implemented and phase 2 was in progress and would be completed in April 2018. The key features of the phase 1, phase 2 and phase 3 development projects are summarised below:

2016 (phase 1):

- Introduced predictive matching algorithm
- Modern, user friendly service web site
- Cleaned up who has access to the service
- Secure code, robust hosting

2018 (phase 2):

- Automated data upload (through API or web upload)
- Transition to XML data schema
- Introduced extended dataset
- Improved data quality and good insight on capabilities of registries

2020 (phase 3):

- Automated patient data upload (through API)
- Development of additional donor matching algorithms
- Automated donor search initiation (through API) - in staging
- Migrated infrastructure to the cloud
- Registry-to-registry communication (through API) - in staging
- Improved data security

The development of a secure registry-to-registry communication system (Connect), or Phase 3, required input from the whole community. WMDA actively promotes participation in the project through Stem Cell Matters (Membership newsletter every three weeks) and during WMDA biannual conference meetings, held virtually since 2020 due to the pandemic. Under the remit of the WMDA's 'Pillar 1' strategic theme, progress reports were presented at the WMDA conference meetings at every opportunity since 2017.

The WMDA community was canvassed via online surveys several times and feedback confirmed a demand to 'democratise' the global donor pool and a desire to improve the functionality of the Search & Match Service. At the WMDA November 2018 meeting, delegates were presented with the latest iterations of the proposed Phase 3 plans (coined Connect), and they provided strong approval to the ability to connect to all registries via a single communication platform.

1.2 Requirements

WMDA strives to significantly reduce the time needed for information processing in the donor/CBU search and request processes and to create fully transparent data exchange - including actual process status - across all registries.

To achieve this, and underpin a safe and sustainable supply of cellular products for its members, WMDA works across three main themes:

1. **Search:** Optimise Search

Ensuring that patients are informed quickly about their chances for transplant is a prime task for WMDA member organisations and transplant centres.

WMDA will actively contribute to this by ensuring:

- All WMDA member organisations are able to search the best available global donor and cord blood data, which is recognised as 'the single version of truth' for all patient searches.
- The proper selection of donors and cord blood products at each moment in time is based on the confidence that donor and cord details presented are up-to-date and checked by WMDA.
- Involving WMDA member organisations where possible. WMDA has and will further strengthen the relationship with and between WMDA member organisations to ensure that good quality data are available in the global file.
- All patients have equal rights to health treatment. WMDA supports this by striving towards a high-quality data source for all patient searches.

2. **Match:** Find the best Match

WMDA identifies best practices in donor search and supports the operations of member organisations. To operate the Search & Match Service a probabilistic matching algorithm will be offered that is accessible in different ways.

WMDA facilitates this by ensuring:

- Search coordinators and transplant centres can register their patients manually, run their search and find the best stem cell source through a variety of filter options in the Search & Match Service.
- Search coordinators and transplant centres can register their patients electronically, run their search and find the best stem cell source through a variety of filter options in the Search & Match Service.
- Search coordinators can automatically run the search from the local software application, receive a match list electronically and use their local settings to select the best stem cell source.
- Search coordinators receive status updates and updates about new potential matches and donor availability.

3. **Connect:** Be connected

Safeguarding and promoting patient and donor care are the ultimate goals of all WMDA members' operations. WMDA will keep in contact with member organisations to ensure that their needs are understood and considered. WMDA members strive collectively to improve the communication between member organisations to ensure transparency and efficiency and to reduce the time to facilitate requests from transplant centres.

WMDA will support this through:

- Accommodating private and secure communication connectors for search coordinators.
- Facilitating easily accessible operational information of member organisations.
- Implementing traceability mechanisms to help search coordinators to keep track of their requests.
- Facilitating low maintenance data exchange for donor/CBU search, testing, typing and work-up requests.
- Allowing registries to make their business decisions and to connect in a well-governed manner.
- Supporting several mechanisms, e.g., web interface and API, to facilitate communication.
- Provide adequate tools for minimising data entry, reducing data errors.

In 2020, WMDA mainly investigated and refined user requirements for the global communication system and worked to ensure security and redundancy of its infrastructure. To learn more about the WMDA's efforts to optimize search, please consult Deliverable *D1.2 Progress report on data quality in EU Member States* for 2020.

Chapter 2. Progress and Achievements, details on how WMDA implemented these requirements are provided.

2. Progress and Achievements

2.1 EMDIS integration into the WMDA

In 1990, the larger European registries developed a peer-to-peer communication protocol using encrypted email called the European Marrow Donor Information System (EMDIS). Starting with only 3 registries connected in this way it grew into a network of 44 listing organisations from all over the globe in the last 20 years. While providing real-time donor and CBU data to one another they also continue to provide their data to the WMDA's Search & Match Service, thereby duplicating this effort. The EMDIS protocol does not make use of open standards in their technology stack and thus creates a high barrier for entry into this sub-community, not to mention the fact that a new organisation would have to make countless peer-to-peer connections compared to organisations that joined in earlier years. Unfortunately, while the EMDIS protocols provide a reliable way to search against and communicate with international donor registries they have been slow to adopt new technologies and security standards.

Since 2019, WMDA has introduced new capabilities through the implementation of APIs that provide all member organisations with secure and controlled transactions between their own and other systems. Initially, the aim of this approach was to provide seamless communication between EMDIS and non-EMDIS registries, with less overhead on setup and maintenance, and to provide a 'translation' layer between different versions of the EMDIS specification. The EMDIS 4.0 Whitepaper - informed by feedback from EMDIS users - signalled an intention to revamp the EMDIS infrastructure, setting the goals to provide modern, commonly accepted technology and greater functionality, that will enable faster communication between EMDIS registries.

However, acknowledging the shared ambition, complementary goals and opportunities for synergy the EMDIS Steering Group proposed the 'integration' of EMDIS into the WMDA. This proposition builds upon a long-standing shared objective to harmonise and align the two initiatives, demonstrated by efforts during 2020 and 2021 to ensure reciprocal representation on steering committees, user and technical groups. Integration would also mitigate the risk that the initiatives are perceived to be in competition or have divergent aspirations and would formalise the collaboration and commonality that characterise the EMDIS and WMDA communities.

On 27 October 2020 the EMDIS community met virtually and by democratic vote approved, in principle, the integration of EMDIS into WMDA on the condition that the EMDIS Chairs provide a clear integration delivery plan by April 2021 that includes the technical blueprint of the combined EMDIS 4.0 and WMDA Connect projects.

A document was developed to serve as framework for the proposed 'integration.' This was a collaborative exercise to appraise the rationale and business case for the potential change, with the output reviewed and approved by the EMDIS community ahead of its endorsement by the WMDA Board.

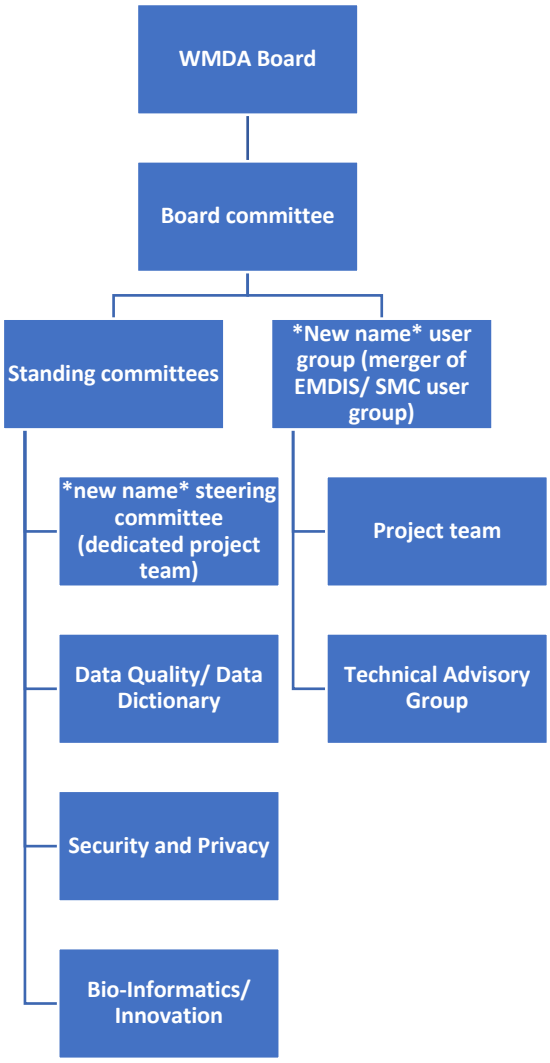
2.1.1 Governance and finance

Principles

- The principle of democratic and inclusive decision-making will remain part of the governance arrangements, with implementing members having responsibility for defining their requirements by setting design principles and parameters. Implementing members will be able to observe design and development and will have an opportunity to test and approve new features/changes to ensure they reflect their agreed upon requirements.

- During the transition period, the Project Delivery Group/Steering Committee shall have autonomy to design and develop products with agility, within the design principles and parameters agreed by users.
- WMDA will provide and/or foster capabilities to enable better communication between registries within the GDPR framework, publish information standards for international exchange of stem cells, and ensure a sustainable operation of the Search & Match Service. WMDA will be responsible for the hosting as well as ongoing support and maintenance of the application.
- The responsibilities regarding data use and ownership have been defined in the data use agreement (which may require amendment). The registries, according to the Data Use Agreement, are data controllers of the donor and patient data. A registry decides if they will set up an agreement with other registries and which/how they will make services available for their international partners. WMDA, according to the Data Use Agreement, is processor of the donor and patient data.

Indicative structure:



Governance during the 'transition phase'

The governance structures of EMDIS and Search, Match & Connect remain basically intact during the transition phase. A special collaborative body ensured the smooth co-development of the project, and communication to each parent organization as well as overseeing the work of the joint project delivery group.

A temporary Project Delivery Group has been set up by the EMDIS and SM&C technical groups for the purpose of defining the outline of the common IT solution and will transform into a Technical Group after the design and implementation of the new common solution on January 1, 2022.

Governance during 'steady state'
After January 1, 2022, EMDIS and Search, Match & Connect will move into the steady state.

User Group

- Responsibilities:
  - Provide insight and expertise to inform the design and development of the single solution.
  - The User Group will write its own house rules, basic framework to be supplied.
  - The User Group directs the development of the application by an RFC process. RFCs having two co-signees from different registries can be proposed by any registry for approval.
  - Test and approve new features/ changes from the Steering Committee's implementation plan of approved RFCs, to ensure they reflect their agreed upon requirements.
  - The User Group will define version support.

- Members:
  - The User Group consists of one voting member per participating registry
  - The assigned voting member will be responsible for casting votes on business and user-based decisions and is therefore responsible for keeping all relevant registry colleagues informed of the current matters.

Technical advisory group

- Responsibilities:
  - The Technical Advisory Group provides a forum for discussion on technical issues of the new solution as they impact on the registries' IT systems and technical workload
  - The Technical Advisory Group provides advise to the User Group on technical questions from the registry perspective
  - The Technical Advisory Group will be self-organising

- Members:
  - All member registry may appoint one member to the Technical Advisory Group

Steering Committee

- Responsibilities:
  - Oversees and monitors the development of the application, ensuring that GDPR is followed, quality of matching ensured, data fields are defined in a correct way and requirements of the implementation users are considered.
  - Reports to the User Group and the Board Committee on progress. Makes recommendations on future direction.
  - Endorses plans and publishes relevant documentation.
  - Maintains oversight of the programme plan, including dependencies and interfaces.
  - Escalates to Search, Match & Connect Board Committee when defined RFCs cannot be achieved within the defined budget.
  - Develops an implementation plan of approved RFCs which will be presented to User Group and who would then have an opportunity to test and approve new features/changes to ensure they reflect agreed requirements.

- Membership:
  - No registry will have more than 2 seats; the chairs of the user group and technical advisory group will serve as members of the Steering Committee.
    - Tech 1: Version control and onboarding
    - Chair Tech 2: IP planning (setting priorities)

- User 1: Receive and process RFCs
- User 2: Testing and deployment
- User Group will appoint a User Group member as its representative on the Steering Committee.
- WMDA office: admin & support

Board Committee

- Responsibilities:
  - Provides overall strategic direction for the initiatives in line with the WMDA's strategy and decisions/views on direction made by the new user group.
  - Is accountable to the WMDA Board for the execution of the WMDA's strategy, as defined and agreed by the WMDA membership.
  - Provides the commitment and involvement necessary to implement change.
  - Establishes and nurtures an environment best suited to achieving programme objectives.
  - Provides updates to members from all work-streams (data dictionary, bioinformatics, data quality, innovation, data security, etc.).
  - Escalates to WMDA Board (when defined RFCs cannot be developed within the defined budget).

- Membership:
  - WMDA Board rep for Pillar 1
  - WMDA Pillar 1 Project Coordinator
  - Chair/member of Security Privacy Committee
  - Chair/member of Data Quality/Dictionary
  - Chair/member of Bioinformatics
  - Chair/ appointed member of new user group

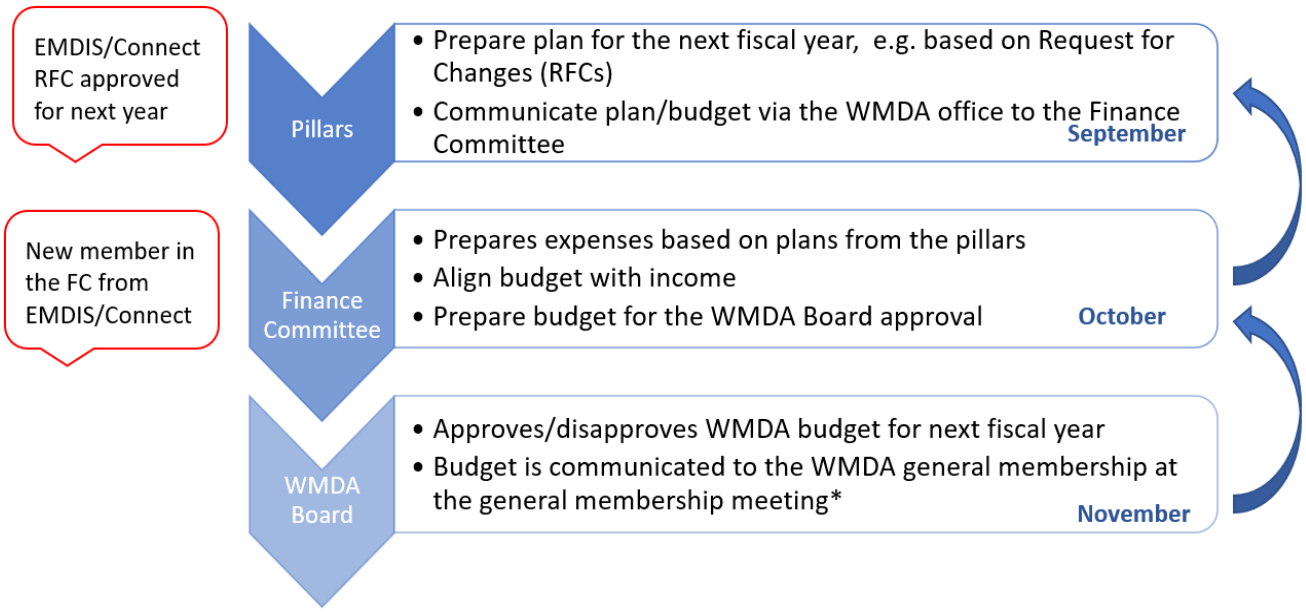
Financial arrangements

The development of the single solution has been covered by the EU development grant and by pro bono contribution of technically proficient individuals from the member organizations. The pace of development of the system is dependent of the availability of funds and human resources.

At steady state the maintenance and development of EMDIS-Connect will be covered by the WMDA membership fee and by pro bono contribution of technically proficient individuals from the member organizations. This will remain, as is currently, the membership fee will reflect the volume of shipments by the registry in question. All members will have access to a core functionality provided by the single solution as part of membership benefits.

Development of additional functionality will be subject to the RFC process and assessment of cost implications by the User Group. If the user group makes decisions with a big cost impact (note: through the RFC process), this will have to discussed and may lead to an increase of the WMDA membership fee.

- In case a substantial increase of the membership fee is needed, the membership will be consulted.
- Registries may choose to develop their own implementations which interphase with EMDIS-Connect, at their own cost
- Monthly finance reports will be sent from the WMDA office to the treasurer, who will update WMDA board.
- The WMDA Finance Committee prepare the budget and annual report, each of which will include a detailed section on the finances of the WMDA software applications including EMDIS-Connect. The WMDA community approves the report and next year's budget. A representative from EMDIS-Connect will be part of the finance committee. The WMDA budget process is included below:



2.1.2 Blueprint group and resulting proof of concept (POC)

In order to develop an integrated technical solution, a working group was formed consisting of technical experts from both WMDA and EMDIS. This working group had a series of 2-hour meetings to compare notes on the EMDIS 4.0 design and the WMDA's approach to registry-to-registry communication.

Herewith a summary of the design decisions made during the blueprint group meetings:

- **XML is proposed for EMDIS 4.0 but switching to JSON used by WMDA would be preferred.**
  - XML was selected by the EMDIS 4.0 pilot because that was the language used by WMDA at the time. Since we haven't started building anything yet, we can still change to JSON.
  - **CONSENSUS REACHED:** All new future communications will be built in JSON

- **AMQP web services**
  - Optional service: Extra work for a small number of organisations that can't open required ports to send and receive messages.
  - Set up a tool for organisations to test whether they have access or make port access a requirement for switch to EMDIS 4.0

- **Could proposed AMQP in EMDIS be switched to API messages?**
  - Yes and no, would depend on the local implementations
  - WMDA landed on API because that is what's used for systems to talk to each other
  - EMDIS 4.0 landed on AMQP to be backwards compatible with older versions of EMDIS that are still using email

- **Proof-of-concept (POC) setup**
  - Show the original EMDIS 4.0 POC in Azure exhibiting document exchange

- **Message queue in this context**
  - Is as at receiving node, they can only read.
  - Important to not develop too many queues to limit the amount of maintenance work required.

- Automation will be essential to assist in communication matrix setup (writing all queues) when entering the environment for the first time.

- **Queue setup (permission structure within queue)**
  - Current EMDIS 4.0 allows everyone to write to every queue, responsibility of the receiving hub to sort out which messages they want to read
  - Communication matrix is unavoidable
  - New registry joins – communication open with everyone, opt out rule for everyone
  - Possible need for a broadcasting queue for global messages

- **What do we do with the unwanted messages?**
  - Automate deletion of unwanted messages on receiver side?
  - Block unsolicited messages on sender side?

- **Encryption critical**
  - However WMDA would need access to decrypted content for translation purposes.
  - Necessary evil to accommodate old protocols.

- **When do messages get encrypted and which keys to use**
  - Does the WMDA play a role in the transit of the messages?
    - Default answer should be no, but:
      - If the messages can stay in the same queue, no decryption by WMDA is required
      - If messages are moved from one queue to another (E3 to E4) via the proxy for the purpose of translation, then yes
      - If messages are moved from one queue to another (API to E4) for the purpose of translation, then yes

- **One-to-one FML vs JSON**
  - Should users know upfront which messages they can send to whom?
    - Expected behaviour = warning message/message denial received from receiving hub (node)
    - Should be able to still read the rest of the message, excluding the new features
  - Two types of messages:
    - Received, can only reply in part
    - Denial, can't reply at all
  - Automated responses from central broker ideal, but this requires decryption of messages by broker
  - Compromise?:
    - Central broker has knowledge about business rules, can deny unwanted messages
    - Central broker sends automated message about receiving registry's capacity to respond
  - Don't have to answer this question right now, but should not develop something without this capability in future
  - Current scenario dictates new IP every year which forces everyone on the same version always, should keep this development schedule

- **Whether it is useful to mark the message type in the AMQP meta information**
  - Useful to include some normalized info for message routing purposes
    - From
    - To

- Msg type
- Version
- Msg format (xml vs fml)

- **Will JSON messages have sequence numbers?**
  - Yes, during transition period but thereafter (when all messages are json) then no longer needed. Sequencing inherent in communication protocol. The central broker assigns an identifier (MessageId) to all brokered messages
  - Sequence numbers might be needed for track and trace of "lost messages or resending specific messages
  - MessageId will change if the message has to be translated by central broker and how would a registry notify WMDA and reference the message when they didn't receive a message?
  - Sequence numbers is therefore needed.

- **Can EMDIScord (CBUs and similar future products) be accommodated by the current design proposal?**
  - Only 6 registries currently make use of it, data mirroring for increased exposure – business benefit
  - Most CBU message types already exist, adult donor cryopreserved unit (ADCU) message types would have to be developed once product category is better defined

Please see Appendix 1 for diagrams and explanations of the POC.

2.1.3 Technical Summit

Unfortunately the progress made by this joint working group was negatively impacted by external parties with divergent goals which meant that several technical questions remained unanswered. In an attempt to answer these questions and also get an opinion on industry best practices, an independent consultancy firm was contracted to engage key stakeholders in a Technical Summit. This summit took place on 6 and 7 July 2021 and was facilitated by solutions architects from Red Badger. Herewith the executive summary from the summit hosts:

Background
The WMDA is looking to modernise and consolidate the systems used by its members and a subset of members within the EMDIS community. Currently the EMDIS members are communicating via an outdated peer-to-peer system which uses email as its transport method. EMDIS is looking to upgrade this system by moving from an email based system onto an AMQP based one. Unfortunately this does not solve many of the problems nested within a peer-to-peer system, these issues are:

- Gaps in the communication matrix
- Increased workload to onboard new registries
- Complicated onboarding process
- Increased workload to update the system leading to infrequent changes and outdated technology

The gaps in the matrix are highlighted by this quote from the EMDIS 4 Whitepaper: "No EMDIS member is connected to all other EMDIS members.". This is especially important and was a sticking point during the summit. Due to the peer-to-peer architecture used by EMDIS any change to the system must be completed by every member rather than once on a centralised service. During the summit this was brought up as an argument for not making changes, despite those changes being the best choice. The diagram below, which was shown in the EMDIS 4 documentation, shows the complexity of a peer-to-peer architecture vs a centralised one.

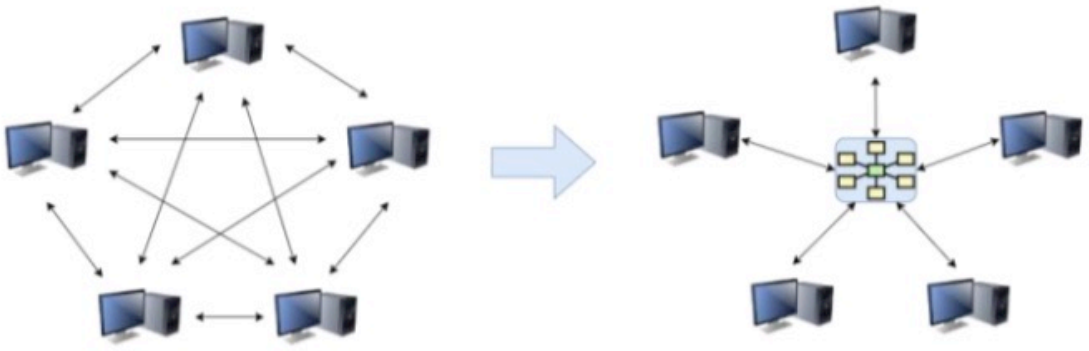


Figure 1: Complexity of a peer-to-peer architecture vs a centralised one

Overall Architecture recommendations

Recommendation: Centralised Service

For these reasons, we suggest any new communications system be a centralised one, with sufficient redundancies built in to emulate the robust nature of EMDIS while reaping the benefits of a centralised service. Those benefits include:

- Single source of truth
- Ability for all registries to connect to each other
- New registries will be instantly connected to previous registries with little effort
- Reduced donor search time
- Reduced change deployment time
- Reduced workload per registry for upgrades
- Traceability, logging and monitoring across the entire system

Part of the hesitation from EMDIS members to adopt a centralised service was that they would lose the availability of the current peer-to-peer system. That is, when one registry goes down the others can still communicate across the network. Currently if one part of the peer-to-peer system goes down it will not hinder the ability of the remaining registries to communicate with each other. In the event that a faulty message gets sent in the current EMDIS system the receiving registry's queue get blocked until it is resolved. In this situation other registries are still able to communicate with one another. In contrast to this, if a centralised service fails then no registries will be able to communicate.

These worries have been mitigated with robust infrastructure design such as using multiple availability zones and auto scaling. Auto scaling can detect when an instance is not responding, stop it, and start a new instance to replace it. Along with best practices such as detailed logging, performance monitoring and an effective rollback policy we think the reliability of the peer-to-peer network can be emulated in a centralised manner. Essentially the trade-off would be between frequent outages in a peer-to-peer system or much less common total outages (when following robust infrastructure design) in a centralised system. It was evident from the summit that the peer-to-peer network used by EMDIS requires a lot of effort to make changes to the system as each registry must update their software. With a centralised system changes only need to happen once, unless the change is to one of the API's, in which case the consumers of that API will need to update. If the API's are managed correctly then these changes can however be taken at a time which suits the registry according to their resources and individual needs. In contrast we're told that EMDIS registries have to coordinate a yearly update between them which leads to a very slow deployment cycle.

An effective web interface should reduce the reliance on registries using the APIs and therefore reduce the changes they need to make to benefit from updates to the system. If registries can complete their main activities via a web interface, such as search and match for example, then they will not need to hook into the APIs and any updates to the web interface will be immediately available to the registries, without them having to upgrade their own systems.

The Connector

While our recommendation is a centralised service, much of the summit discussion related to incorporating a form of connector in a final solution. The idea being that the connector will link the centralised WMDA solution with the peer-to-peer EMDIS solution. The discussions during the summit were as follows:

- Vendor choice
- Sending large documents
- EMDIS 4 document a 'spec'
- Encryption vs Translation

Each of the topics were discussed in full and Red Badger's opinion and best practices guidelines given.

Sending Large Documents

What is it?

In the future both WMDA and EMDIS systems will need the ability to transfer documents between registries. This discussion captured the best ways to do this.

Capturing the discussion

- EMDIS often sends messages of 20MB+ daily
    - Often these are multiple messages bundles together
- Azure can't send messages over 1MB
- Other message brokers can send messages larger than 1MB
    - Most of these will need the default settings changed to do so
- Some EMDIS implementations already have the capability to split messages
- Authentication configuration may need to be duplicated across two services, a message broker and a file storage system

Discussing solutions

External file storage (suggested approach)

Large documents can be stored in an external service and only the storage location is sent in the message. This is known as the claim-check pattern. This follows industry best practice of keeping messages small in size and therefore helping to alleviate some of the problems that come with larger messages.

Increasing size limit

This would need the fewest changes from an EMDIS perspective but would require a change of infrastructure for the WMDA, this is because Azure currently has an upper limit of 1MB for messages. AMQP is designed to send small messages and can suffer from performance issues when messages are too large. This is why it is recommended to keep messages below 100kb.

Splitting messages

While message splitting can make messages smaller it is much riskier. If a message producer were to fail while sending a split file then the consumers will not be able to reproduce the completed file. A solution

to this would need to be implemented in either the producer, consumer or both and would be quite complex. Splitting files into multiple messages would also prevent the messages from being atomic.

Best practice in this area

Industry best practice is to keep messages small in size, helping to alleviate some of the problems that come with larger messages. One reason for issues is that the message has to be fully processed by the broker before it is made available in a queue, this requires more resources on the broker. Additionally, the whole message must fit into memory on the broker, which wouldn't be a problem for a single message but could cause issues when scaled to many registries all communicating in tandem. These problems can include long processing times, high ram usage and backed up queues.

It is also generally best practice to keep messages as atomic as possible when using a message bus. This is because messages can often arrive out of order when using a message bus, this is especially true if there are multiple producers/consumers to a queue, which there may be in the future. Atomic messages also allow for multiple consumers to split the workload as each message can be processed independently and does not need prior knowledge of state that may be stored on another machine. The ability to balance the workload across multiple machines in this manner allows the system to scale without losing performance.

Red Badger recommendation

Our recommendation is to store the file externally to the message bus and send a reference to it in the message, this is otherwise known as the claim-check pattern. This allows the WMDA to use a service purpose built for storing and serving files such as AWS S3 or Azure Blob, which would result in lower costs than storing the file on a message bus.

We would also recommend using content addressable naming conventions to ensure the specificity of the reference, this would likely result in the reference being a sha256 hash of the file being sent.

Outbox Routing

What is it?

AMQP uses the concept of queues to organise and deliver messages, this discussion focuses on the read/write strategy surrounding these queues. Two queue models have been proposed, one is to have each registry produce messages directly to the queue of the intended registry. The other is to produce the message to their own topic and then have a central routing service forward it to the correct queue.

Capturing the discussion

- Authentication complexity is increased with every new registry without an outbox
    - This can be avoided on some platforms
- An outbox would need a routing element
- Message routing is a widely solved problem
- A message router would introduce a new point of failure

Discussing solutions

Outbox forwarding

The outbox model would require less configuration when managing keys as only one write permission would be needed per registry rather than the peer-to-peer setup of having each registry require write permissions for each registry they wish to communicate with. This would reduce the need to set up additional permissions for every registry when a new one joins. While Azure and AMQP (and by extension AWS) have the functionality to forward messages either built in or via a plugin, it would still add a small amount of complexity to the solution.

This however does seem small. Adding headers to the messages should be a minimal amount of work and would allow for routing to be configured.

Producing directly to the recipient's queue
Producing a message directly to the recipient's queue would require more configuration and would not allow for any centralised pre-processing validation (depending on whether the messages are end-to-end encrypted or not). The lack of centralised message validation would mean that each registry would need to validate the message themselves and the level of validation could vary between registries. An invalid message that is not correctly filtered out could cause issues for that registry.

By not including a centralised message router you reduce the slight risk of this message router becoming unavailable and bringing down the wider system. Although we do think message routing is a simple and widely solved task and fairly low risk so providing high availability should not be an issue. There is also a possibility that if the message router does not have sufficient resources then it could cause a bottleneck within the network, this however could be mitigated easily by following best practices and implementing auto scaling.

Best practice in this area
Generally it is best practice to validate messages before allowing them to leave the system to be consumed by an external party such as a registry. Using an outbox model would allow for this. It would also be more scalable as the authentication configuration would be significantly reduced.

Red Badger recommendation
We would recommend using the outbox model.

Azure vs AWS
What is it?
Azure and AWS are both very widely used and trusted cloud hosting platforms. Azure is run by Microsoft and AWS is run by Amazon.

Capturing the discussion
- Both Azure and AWS are permitted for use in most government institutions
- The WMDA team are more experienced with Azure than they are AWS
- Using Azure will require more changes for EMDIS members due to platform limits

Discussing solutions
To use Azure the current EMDIS software must be updated significantly to follow industry best practices. This will increase the EMDIS workload considerably. Inverse to this, using AWS will be new to the WMDA and will require them to either manage two platforms or perform a migration of their existing systems to AWS.

Best practice in this area
Given best practices are followed both AWS and Azure offer largely equivalent services and tools and should be capable of hosting the connector. During the summit the issues mainly seemed to revolve around work needed for either party to switch infrastructure providers. Either the EMDIS community must spend time adopting best practices in order to use Azure or the WMDA engineers must spend time learning and migrating it's infrastructure to AWS. This must be discussed by the two parties and comes down to more of a cultural decision.

Red Badger recommendation
We would advise that best practices are followed when building any system, in which case Azure or AWS can be used. The decision on a platform then should be based on multiple other factors such as cost, ease of use, maintainability and performance.

2.1.4 Match-Connect

In consultation with the WMDA Finance Committee, the WMDA Board approved the integrated governance structure and technical recommendations made by Red Badger. Based on these two documents a Steering Committee was elected to drive development of a communication solution described by industry best practices. The Steering Committee will propose the following options to WMDA members:

Option 1 (WMDA API series):
Adopt a new streamlined data-interface technology for exchanging information among registries

Option 2 (Adapted EMDIS 4.0):
Adopt the updated EMDIS Communication System and protocols to ensure that they communicate with an infrastructure that facilitates communication between EMDIS registries and between EMDIS and non-EMDIS registries

Option 3 (No action required):

- If you are not on the EMDIS protocol, you can continue to use manual processes like fax and email.
- If you have the EMDIS protocols implemented, you can continue to use the current EMDIS 3 as it will continue to work, however support for it will be discontinued in future and no new features will be developed.

Option 1 is the recommended choice towards a single communication solution. Option 2 is offered to ease the transition from EMDIS protocols to the APIs and will be phased out over time. Option 3 is not recommended but an option required for organisations that are based in countries that prohibit the communication of personal information using this technology. These options have been developed and offered to the WMDA members in December 2021.

2.2 Alternative matching algorithms

Historically, the Search & Match Service uses the Optimatch matching algorithm which provides probability matching using haplotype frequencies calculated using the entire global donor and CBU database as a dataset. Unfortunately, due to licencing constraints, the Optimatch algorithm could not be implemented into the Search API and alternative matching algorithms had to be investigated. Algorithms from two large member organisations were considered, namely ATLAS from Anthony Nolan and NMDP jointly, and HAP-E from DKMS.

Please see Appendix 2 for the slides presented to the community during an educational webinar.

2.2.1 HAP-E

The HAP-E Search algorithm has been in use in the DKMS internal donor search system since 2011 and was further optimized in 2018 by adding the ability to perform genotype recalculations (reference: "Urban, C., Schmidt, A.H., & Hofmann, J.A. (2020). Hap-E Search 2.0: improving the performance of a probabilistic donor-recipient matching algorithm based on haplotype frequencies. Frontiers in medicine, 7, 32."). This reinforces confidence and validity in its calculations as they have been verified repeatedly. However, development is required to scale it up to make it robust and accurate when applied to the much larger and more diverse global donor data pool took time. HAP-E is now offered to search coordinators on a global scale.

2.2.2 ATLAS

In 2020, Anthony Nolan launched a brand-new search algorithm for internal use, developed in collaboration with Softwire and NMDP. The aim was to build a world-class search algorithm, incorporating the latest research in HLA matching, and leveraging the power and flexibility of cloud computing. Project ATLAS is a continuation of that work: enhancing the algorithm to make it suitable for the widest possible audience, for the benefit of the global community. Atlas is offered as an open-source solution, so that all registries in the world can benefit from it including WMDA. It gives the opportunity to validate the search, because the search coordinators are able to do a search with HAP-E as well as with Atlas.

2.3 Data Security and Privacy

The WMDA's Security Privacy Committee encourages and enables the WMDA and its members in managing information security and privacy matters as a global business risk. With proper risk management, WMDA can ensure the long-term sustainability of global data sharing that is necessary to achieve its mission.

The WMDA Security and Privacy Committee (WMDA SPC) achieves this by:

- Establishing minimum security and privacy standards that are required for WMDA membership.
- Supporting the WMDA Accreditation process to ensure these standards are enforced.
- Providing regular education and training to members at semi-annual meetings and with publications.
- Informing members on emerging risks and regulations that may threaten the WMDA mission.
- Providing practical guidance specifically crafted for small- and mid-sized stem cell registries with an international supply chain and data profile.

2.3.1 Penetration testing

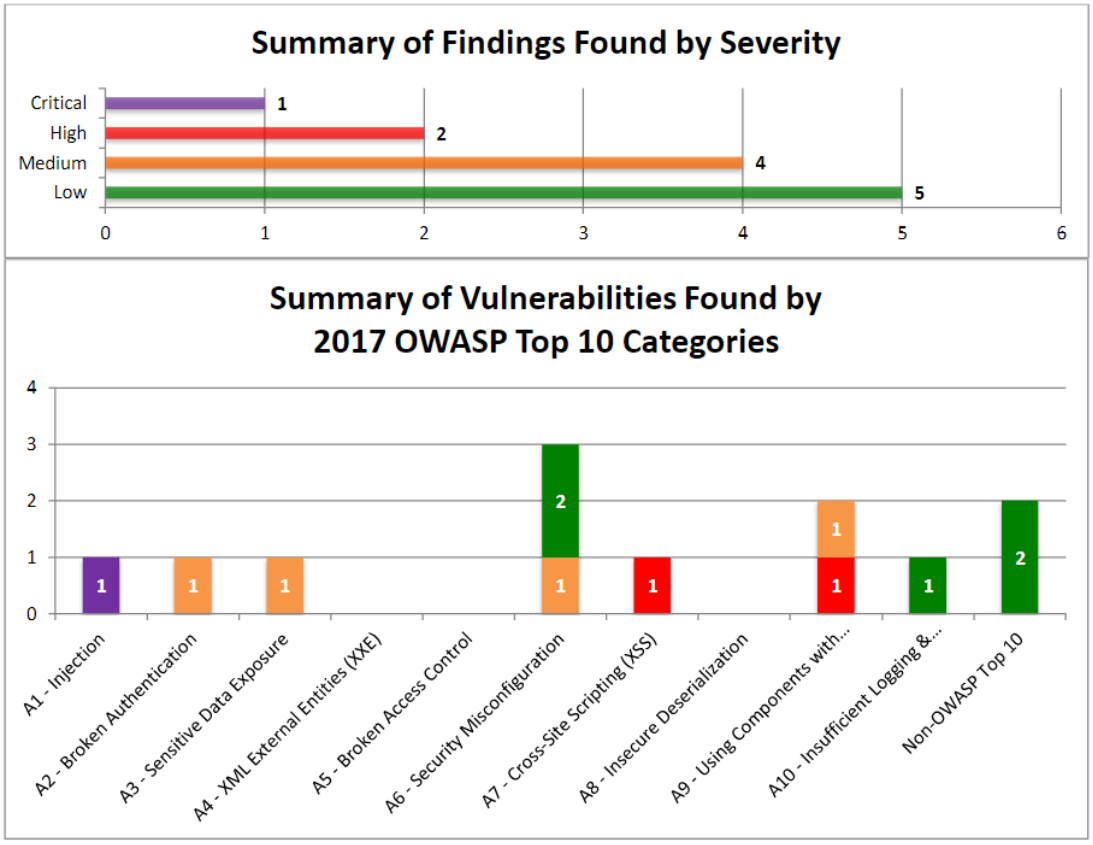
Between December 16, 2020 and December 29, 2020, Be The Match engaged White Oak Security to perform a Web Application Penetration Test of the WMDA's Search & Match Service and Data Upload applications. The goal of the engagement was to determine if sensitive customer data could be accessed by unauthorized

individuals and to ensure that the Open Web Application Security Project (OWASP) Top 10 were being adhered to by the application development teams.

The web application penetration test discovered one critical-risk and two high-risk issues. The first two issues affected Data Upload exclusively. The remaining findings were medium-to low-risk configuration and patching issues.

During the engagement, White Oak Security identified twelve (12)findings in total including:

- one **(1)Critical-Risk issue**,
- two **(2)High-Risk issues**,
- four **(4)Medium-Risk issues**, and
- five **(5)Low-Risk issues**.



All the security vulnerabilities identified were addressed accordingly by eliminating or mitigating the risk. While thorough security monitoring will reduce the risk of data loss/theft it comes at an equally high cost. WMDA would need to contract a professional organization for that. First estimates from external organisations range from 60.000 to 90.000 Euro annually when all is covered, excluding the extra setup costs. Currently the WMDA is only working on getting the proper monitoring in place for Azure, having periodic vulnerability checks performed and implementing ISO-27001 related procedures within its own systems.

In view of the GDPR implementation in the Netherlands, it's worthwhile mentioning that the executive director of WMDA is personally liable if no proper action (negligence) is taken with regards to information security. Should the members opt against the implementation of sufficiently strict rules in WMDA, this liability will have to be taken over by the members as it cannot be insured.

2.3.2 Multifactor Authentication

As part of the WMDA's efforts towards risk mitigation, a multifactor authentication (MFA) solution has been developed and will be rolled out to the global membership to authenticate user access to the data held by WMDA. The services of a solutions architect were contracted to perform the investigations, develop a proof of concept and design the roll-out. Following investigations into some "boxed" solutions like OKTA and Auth0, it was decided to consider a more affordable "in-house" approach by leveraging the Microsoft Azure cloud services WMDA already had access to since migrating its infrastructure to the cloud (See Deliverable *D1.2 Progress report on the implementation of a secure registry-to-registry communication system – 2020* for more information).

Azure Active Directory (in short – Azure AD) is a cloud identity provider service or Identity as a Service (IdaaS) provided by Microsoft. Its primary purpose is to provide authentication and authorization for applications in the cloud (SaaS apps). Azure AD's main purpose is supporting business organizations with extending their identity reach to the cloud and Software as a Service (SaaS) applications. On top of that, there are tons of enhancements and services provided, such as conditional access, identity protection, application publishing, access to pre-configured applications etc. which allows WMDA developers to build applications and secure them with Azure AD. However, to make these applications available to our membership we need to extend the authentication to our member organisations using Azure Business-to-Business AD (B2B) and Azure Business-to-Consumer AD (B2C).

By leveraging these service, WMDA:

- Saves on future licensing fees inherent to "boxed" solutions
- Can utilize the benefits of both B2B and B2C service making it more user friendly and easier to onboard members
- APIs can seamlessly implement the same authentication solution as they are built in the same environment.

The high-level requirements are described as follow:

Must-haves:

- Centralised, so MFA is not handled in the connecting application
- Separate identity store. E.g. Azure AD B2C instead of directly in Azure AD.
- MFA must work all over the world
    - Choose different second factor for people in different organisations
        - e.g. force use of a strong MFA methods (hardware tokens, OTP apps etc.) by default, but allow exceptions for countries where this is not allowed.
    - Some countries are more of a challenge because they may not support the standard MFA methods such as Google Authenticator. Therefore the following countries require extra attention to check whether there are supported authentication methods:
        - China
        - Iran
        - Saudi Arabia
        - Russia
- 400-500 users are expected to be needed
- Current CRM stays the "source of truth". Some options are:
    - Every x minutes, the MFA solution imports de users AND their authorisations from CRM. For example using an XML, as is currently also used in Search and Data upload. Updates of users and their authorisations do not happen often, so synchronisation would not need to happen very often. Currently it is every 2 hours, which is fine.
    - The identity solution directly checks with the CRM whether user is authorised to access the application and if so, which access rights the user has. The WMDA CRM is currently considered a

"zone 2" application and can therefore be down for a longer period that e.g. Search. Therefore this option is not preferred, because when CRM would go down, nobody would be able to login any more.

- The following applications should be able to connect without much work on the MFA solution.
  - Search.
  - Data upload.
  - SPEAR rebuild. This application is being rebuilt at the moment. A connection to Azure AAD/ Azure B2C is being kept into account.
- The following applications should be able to connect to the MFA solution. It is OK if some work is needed on the MFA solution in order to make it connectable to the applications described below:
  - Atlassian Confluence ("WMDA Share")
  - Moodle ("WMDA Education"). Should be able to use Azure AD connection. Possibly also other connections.
  - Espocrm ("CRM")
- Password policies:
  - Force change of password at a set interval
  - Password complexity requirements

Should have:

- A user has same set of credentials for all applications connected to the MFA solution

Nice to have:

- Connection between MFA solution and Azure AD so that WMDA employees can use their Azure AD credentials for logging into applications. This eases the process of logging in for WMDA Azure AD users.
- Single sign on (SSO) between applications, even if different authentication methods are used.

Should NOT have:

- Authentication via social media connections. Would bring up all kinds of questions regarding security/privacy. For now just consider SMS, phone, email, common authenticator apps (Google, Microsoft, Lastpass e.d.) possibly fido-keys

At the time of writing, both the B2B and B2C MFA solutions had already been tested for implementation on all the APIs and were in the progress of testing with members from countries identified as possibly problematic.

2.3.3 Peer security assessment

WMDA Security and Privacy Committee (SPC) established a simple, scalable, and repeatable assessment process to perform registry-to-registry peer risk assessments of registry applications. One security, privacy, or IT representative from WMDA performs a security and privacy assessment on behalf of the entire WMDA network. Basic information on the results, and any material corrective actions, are published on WMDA Share. Members may rely on these results to comply with their own internal and external risk assessment requirements. Large registries and the WMDA are the early adopters of this new process.

This process was piloted between NMDP (USA) and ZKRD (Germany). Thereafter, the following peer reviews took place:

- Matchis Foundation (Netherlands)
- Australian Bone Marrow Donor Registry
- Anthony Nolan (UK)
- Redome (Brazil)
- DKMS (Germany) – pending

- Ezer Mizion (Israel) – pending
- China Marrow Donor Program – pending

2.4 Refactoring internal infrastructure to accommodate technology stack changes

As a non-profit organisation and global association, the WMDA office consists of a small team of employees but has access to a wealth of expertise in its membership. The ongoing collaboration between the WMDA office team and international partners is mutually beneficial and has led to many successful projects. The downside however is that on long, multi-year projects there is a lack of continuity as people with relevant expertise moves in and out of the community. This results in fluctuations in the momentum of the project and slight misalignment between different elements of the project.

Since the original conceptualisation of WMDA's vision for registry-to-registry communication in 2018 up to now, the project has benefitted from the involvement of close to 100 experts, consultants, project managers, working group, steering committee and board members, not to mention developers and 3rd party contractors. While this has brought us very close to achieving our goal of a single registry-to-registry communication platform it has also left us with what can be described as a hodgepodge code base. In order to ensure that this solution is future proof, we decided to take the time and invest the effort to align the code base and establish clear design principals.

2.4.1 Security improvements in Azure

A Web Application Firewall (WAF) is a firewall that is specifically designed to scan web traffic. Where a traditional firewall is only able to allow or deny traffic based on IP address and or port, a WAF can actually look inside the traffic flows and detect malicious behaviour. Some of the main types of attacks detected and blocked by WAF's are cross-site-scripting (XSS), cross-site forgery, file inclusion, SQL injection. It acts as a shield between the internet and the actual web server, stopping attacks that would otherwise have successfully made use of a security vulnerability in a web service.

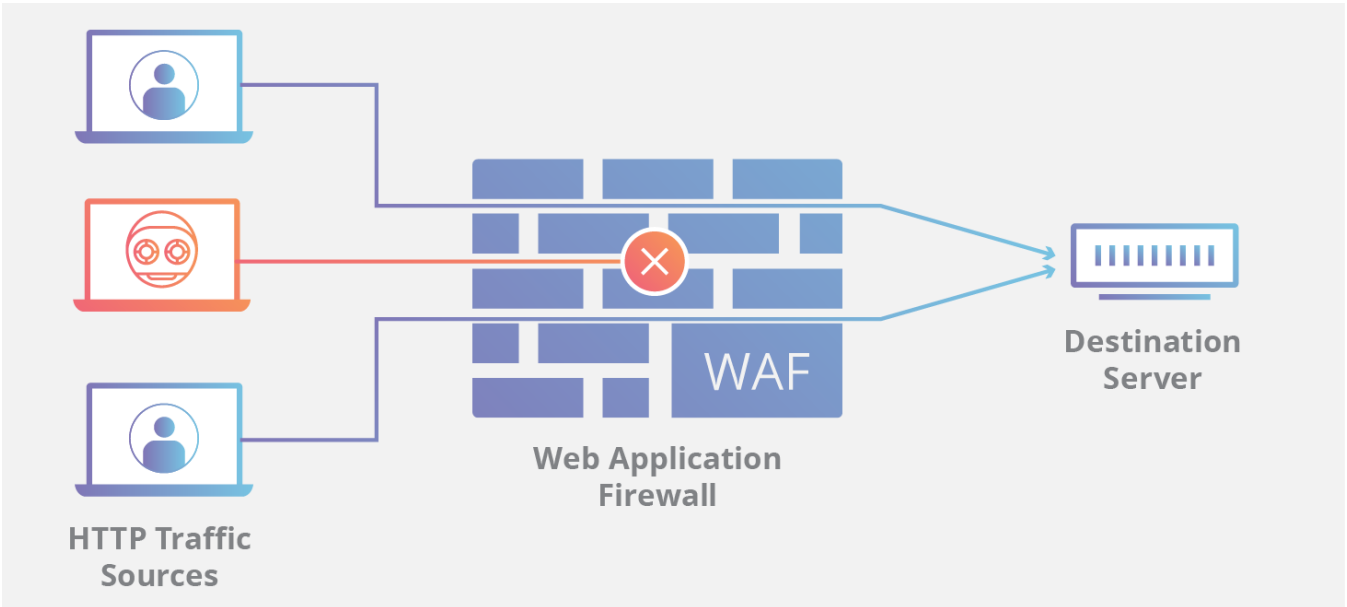


Figure 2: Visual representation of the function of a WAF

A WAF uses combinations of rules called policies. In the case of WMDA one of those policies serve to block known malicious bots. This is based on a list that is maintained by Microsoft. Using this list, known malicious bots are not able to even connect to one of WMDA's web services because it is "blocked at the gate". Another

policy the WMDA uses is based on the Open Web Application Security Project (OWASP) ModSecurity Core Rule Set. This is a set of generic attack detection rules maintained by an open-source community.

This OWASP ModSecurity Core Rule Set does not have rules that are specific for a certain attack variant, but instead the rules are more heuristic in nature. They are able to detect malicious patterns. The upside of that is they can protect against not-yet-known types of attack that still follow a malicious pattern, but the downside is that it may occasionally label legitimate traffic as malicious and blocks it. Because of this, each application that you want to run through a WAF will first need to go through a period of time where the WAF is not actually blocking the traffic but merely marking it as malicious. WMDA Azure administrators can then monitor the traffic and detect false positives without the WAF interfering with legitimate traffic. If necessary, certain rules within the rule set can then be disabled for an application because the rule was triggered for legitimate traffic. In some cases the application can be adapted to no longer behave in this suspicious manner. After this adaptation, the rule can be enabled again, but in some cases this is not possible, because it concerns an off-the-shelf application that cannot be modified in that way.

When it is found that no legitimate traffic is marked as malicious any more, the WAF can be switched to "prevention mode" where it will actually start blocking traffic marked as malicious. The process of screening traffic, tweaking rules and then fully turning on the WAF for an application is one that takes time, but is worth the effort because it tweaks the WAF settings to the specific application and therefore provides optimal security.

In 2021 all services hosted by WMDA were placed behind this combination of Application Gateway with Web Application Firewall, significantly enhancing security. Where possible, hosted applications were adapted to optimally run in this environment.

2.4.2 Team evolution

The WMDA's hodgepodge code base consisted of bits of Java, PHP and C# in various styles. With the migration of the internal infrastructure to the Azure cloud environment, it was decided to conform all code to C# in a single code design and branching strategy. To do so we required developers experienced in multiple coding languages and styles.

January 2020		December 2021
2 x Product Owners 1 x IT consultant 1 x Senior developer – external 1 x Project manager – external		1 x Business Owner 1 x IT Consultant 2 x Product Owners 3 x Senior developers 1 x Junior developer 2 x Intern developers 1 x Solutions Architect – external 1 x Database Administrator – external 1 x Project manager – external

The large addition of experienced staff to the internal team meant a change in the office space, team dynamics and salary budget of the WMDA. The international scope of the projects is reflected in a diverse team that boasts male and female members from Peru, China, India, Spain, South Africa, UK and of course the Netherlands. The WMDA leverages expertise from its membership but takes a collaborative approach rather than an "outsourced project" approach to ensure that the knowledge remains in the team.

2.4.3 Agile development methodology

With the evolution of our internal IT team, the way we worked also had to evolve. The Agile Methodology, popularised in 2001, was adopted and over time applied in such a way to suit our ever-growing team.
In software development, Agile practices include requirements discovery and solutions improvement through the collaborative effort of self-organizing and cross-functional teams with their customer(s)/end user(s), adaptive planning, evolutionary development, early delivery, continual improvement, and flexible responses to changes in requirements, capacity, and understanding of the problems to be solved.

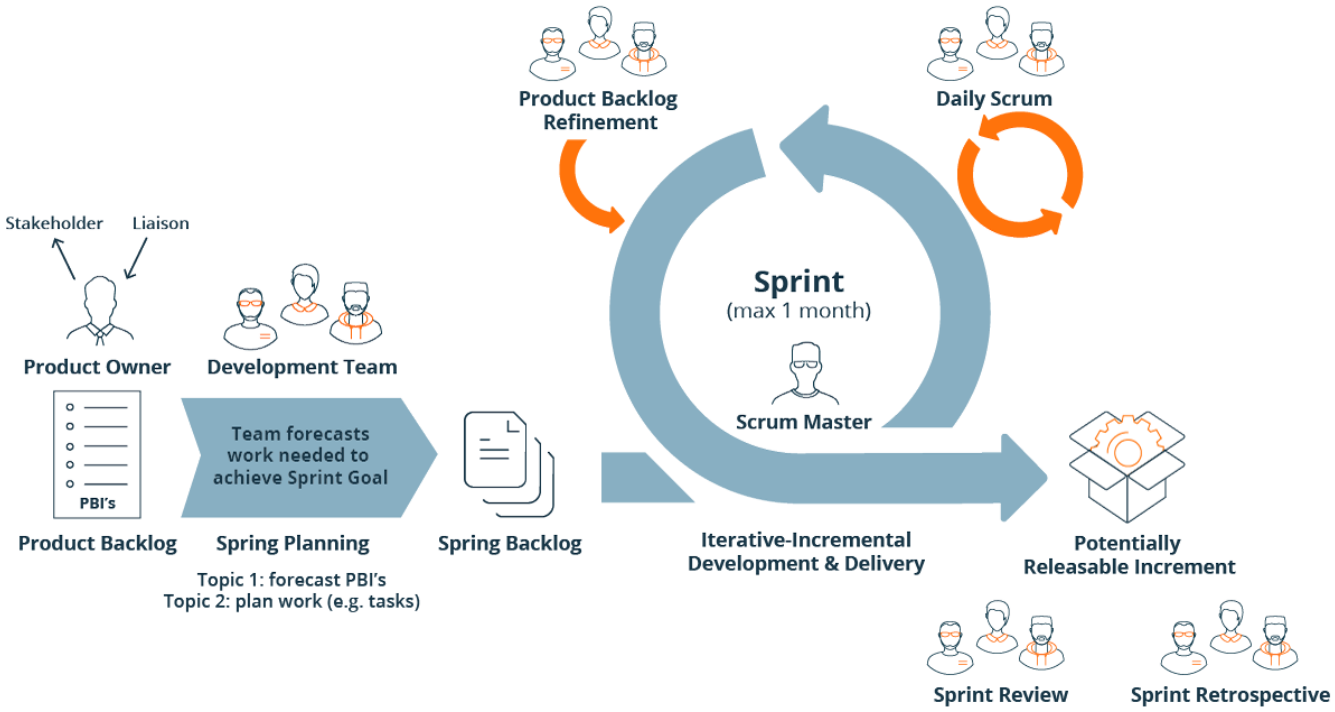


Figure 3: Agile development life cycle

For every internal IT project the project scope and requirements are defined prior to defining the roles and responsibilities of the team members involved, as members might fulfil different roles on different projects. A sprint cycle generally lasts 2 weeks, begins with a refinement meeting and ends with a cycle retrospective/demo and backlog meeting. Short daily stand-up meetings ensure that the project stays within the scope and on time by addressing blockers as soon as they present themselves. This process has proved to work well in scenarios where international partners and 3rd party consultants are required to join or leave midway through a project. The WMDA benefits from the external input while maintaining the knowledge internally.

In Chapter 3 a high-level strategy for future development, as well as support and maintenance of the existing solution is presented.

3. Beyond 2021

3.1 Future development

The technical specifications how organisations can easily communicate with other organisations through a single connection are available. The Connect API developed in 2021, has enabled communication between non-EMDIS registries as originally envisaged by WMDA. The integration of the EMDIS community and the melding of the two solutions and communities has created opportunities.

The APIs and EMDIS 4.0 will be extended to support each other, provide additional processes, status updates and auto-complete forms. WMDA will endeavour to implement a secure registry-to-registry communication solution based on the original scope of the Connect API in collaboration with the EMDIS Technical User Group and their vision of EMDIS 4.0. The integrated solution (EMDIS-Connect) will be delivered step-by-step ('agile' delivery) and will enable registries to connect their internal system to the global community allowing Search Coordinators to:

- Send and receive messages between non-EMDIS and EMDIS users based on EMDIS protocols.
- Perform donor availability checks, subject to data quality dependencies.
- Conduct donor health and availability requests, extended HLA and HLA verification typing requests, donor infectious disease marker requests, donor workup requests and CBU shipment requests.
- Receive automated acknowledgement, receipt and status notifications.

From an IT management point-of-view the following core principles are safeguarded:

- Low disruption to existing systems and processes and avoidance of 'forced systems migration'.
- Ease of implementation to establish connections to all registries.
- Low maintenance and overhead for registries.
- Security and data protection 'built in' to user authentication and message handling solutions.
- Single data model based on EMDIS semantics and used by all registries.
- Data quality and governance steered by WMDA and supported by the Data Dictionary Working Group.
- Improved transparency and reporting of data flows.
- Centralised notification and support for process and business rule related issues (e.g., invoicing problems)

The WMDA's vision will be achieved by realising the following goals:

- One single Search, Match & Connect data exchange hub (described as WMDA Connect in figure 1)
    - Facilitate effective and efficient communication between registries.
    - Facilitating registries by offering one single connection point to connect to, for communication to all registries.
    - Translation of messages in different data formats including EMDIS.
    - Based on EMDIS data semantics; and keep an eye on development in HL7/FHIR.
- Implementation of APIs to Search, Match & Connect for real-time data uploading downloading (described as API in figure 1)
    - Search coordinators can obtain the same information in their local system through an API as they can obtain when they register a search online through the web application.
    - Development of a secure, robust, simple and scalable API.
    - Non-disruptive, backward compatible connector, that allows registries to continue to use their current operational registry software application.
    - Compliance with European data regulations (GDPR), reducing the need to create a data hub or to have a
      downloaded copy of the dataset.
- Extended functionality connectors with respect to EMDIS (described as API in figure 1)

- The connector provides the possibility to exchange more data including documents.
- The connector supports more processes (e.g., verification and extended typing, infectious disease marker
  testing, work-up requests)
- The connector provides status updates, e.g., request received, request read, etc.
- The connector provides automatically filled-in forms.

- Maintaining master data on behalf of registries (described as API in figure 1)
  - The connector will allow registries to define their preferences to decide how to collaborate with their international partners.
- Improved functionality (described as API in figure 1)
  - Storing/re-use of search criteria (profiles).
  - Define user/transplant centre preferences.
  - Quality monitoring.
- Use of modern technology
  - Low maintenance, non-specialized ICT team available once implemented.
  - Widely used technologies in-line with standard industry practice, preferably open source.
  - Implemented via an agile framework, step-by-step delivery to the WMDA member registries.
  - Cloud native approach allowing for distributed processing, improved resilience and optimized performance.
  - Open, allowing for inclusion of community developments (added functionality).
  - Based on “secure by design” and “privacy by default” with full traceability of performed actions.
  - Use of automatic quality checks to reduce test effort and improve quality of delivery.
- Improved availability of Search, Match & Connect
  - Improved hosting environment: Connect requires high availability across the world and a scalable
    environment which is maintained 24x7; business requirements will be developed in collaboration with
    the Search, Match & Connect Steering Committee.
  - Continuous delivery approach for functional enhancements.

By combining the existing EMDIS proxy with the EMDIS-Connect communication solution, secure registry-to-registry communication will be available to all WMDA member registries and will pave the way for complete integration of EMDIS into the WMDA.

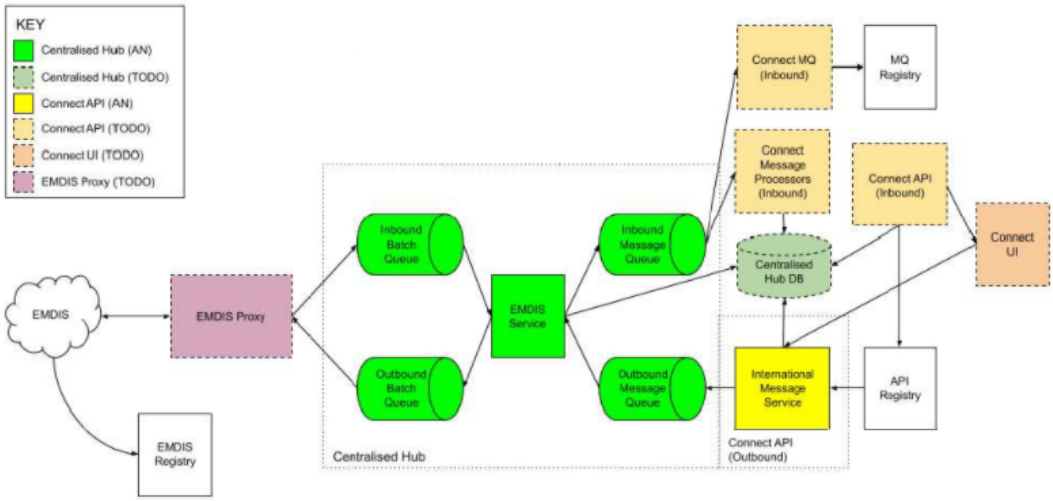


Figure 4: Proposed technical overview of communication flow between EMDIS & non-EMDIS registries

3.2 Support and maintenance

In consultation with the WMDA Finance Committee, the WMDA Board approved the integrated governance structure and technical recommendations made by Red Badger. Based on these two documents a Steering Committee was elected to drive development of a communication solution described by industry best practices. The Steering Committee has described the technical specifications (see appendix 1) of the communication solution which will be hosted and maintained by the WMDA. An addition to being an accreditation body, educational platform and global authority on donor care, WMDA will now also provide software as a service to its member organisations.

This will require a service level agreement between the WMDA and the member organisations that implement the solution that defines the terms of the service provided. Additional resources will also be required in the form of data logging and message tracking for trouble shooting purposes, a ticketing system to report errors, first, second and third line support to resolve tickets and monitoring the general health of the central broker.

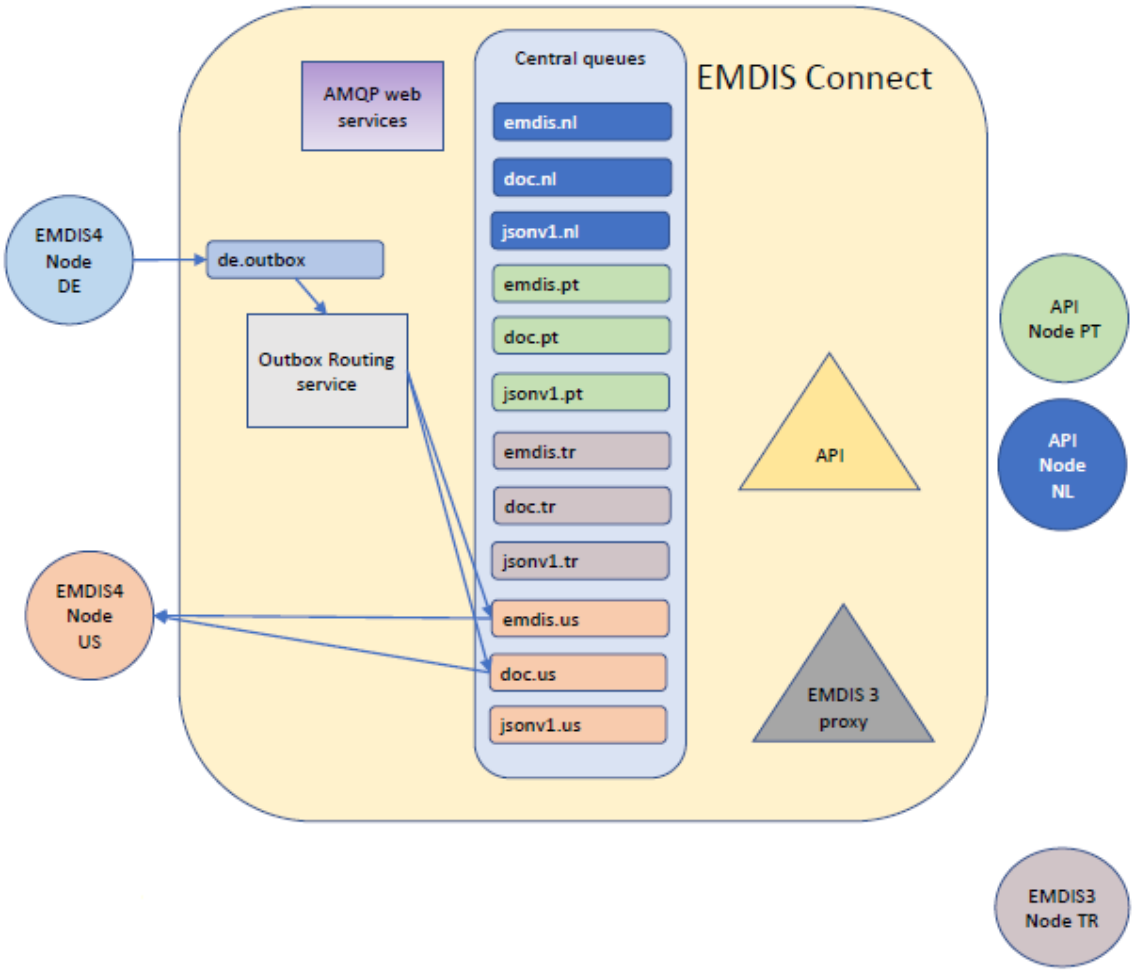
This industry is dynamic, ever changing and needs to be at the frontlines of medical and scientific advancement to ensure quality patient and donor care. This is only possible if the communication infrastructure can meet those demands therefore development cannot stop or it will stagnate. As a community driver association, WMDA will rely on its members to drive the direction of future development through an agreed upon procedure to request changes and updates to the technology.

Request for Change (RFC) Procedures
The User Group directs the development of the application by an RFC process. The members will request developments or revisions using a standard format for RFCs, as determined by the members of the User Group.

- RFCs shall be submitted to the Chair of the User Group using the RFC submission form.
- An RFC shall be supported by two other members
- The User Group Chair (and/or Chair-elect) will review the request for completeness and evaluate whether it reflects user requirements.
- Requests that fulfil these requirements will be passed on to the Steering Committee.
- The request will be reviewed by the Steering Committee, and if needed, the Steering Committee will assign a technical expert from the Steering Committee or Technical Advisory Group to work with the original requester to elaborate on the technical requirements.
- All completed RFCs will periodically be presented to the User Group for final approval and prioritisation.
- An RFC requires a two-thirds majority of the delegates present at the meeting to be accepted.
- The Steering Committee will develop an implementation plan of approved RFCs.
- Any issues (e.g. financial) with RFCs discovered during specification in the Steering Committee are reported back to the User Group to be reviewed.

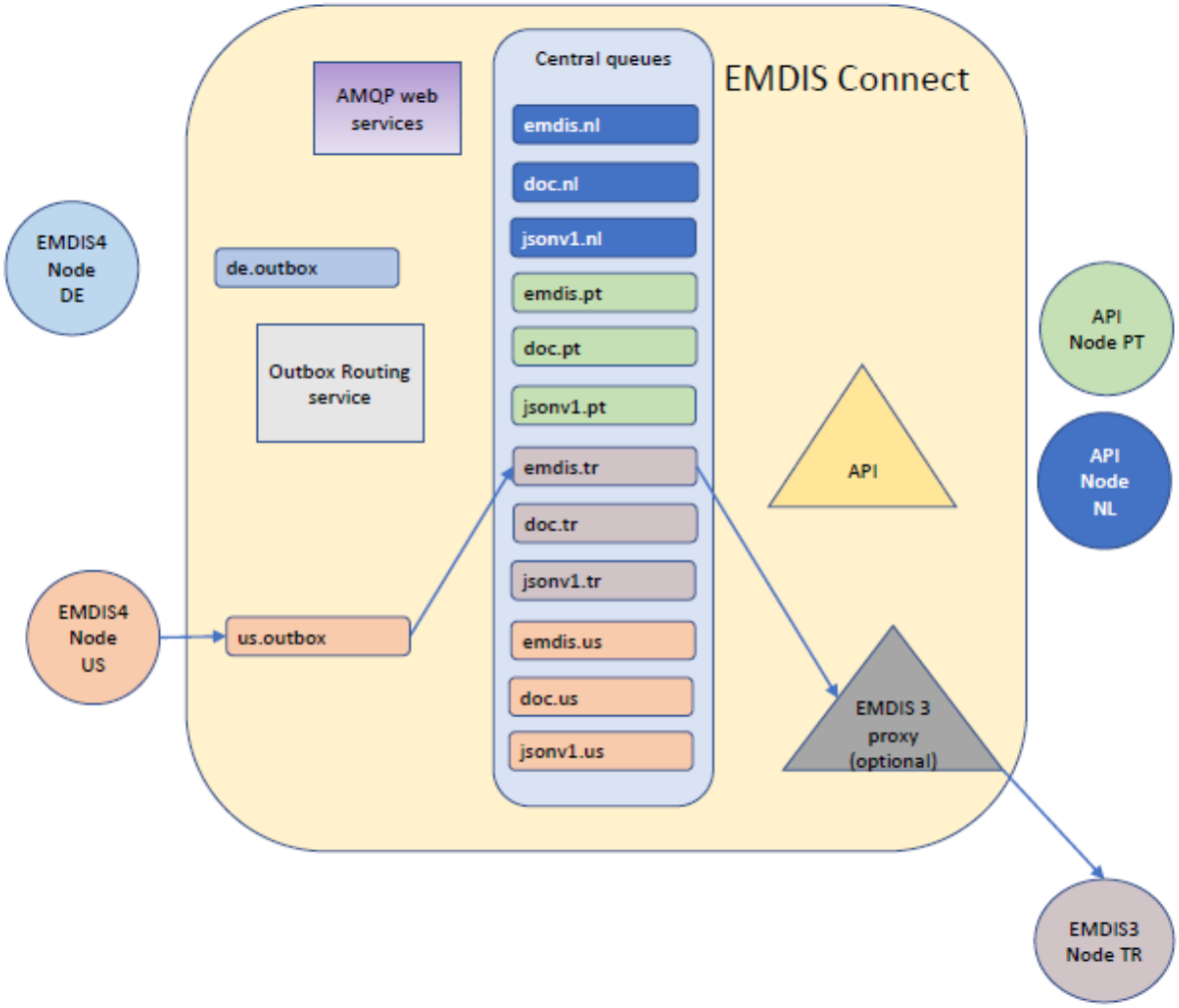
Appendix 1 – EMDIS 4.0 and WMDA Technical Integration – POC design

From: EMDIS4 (DE)
To: EMDIS4 (US)
WITH "inbox" & "outbox"



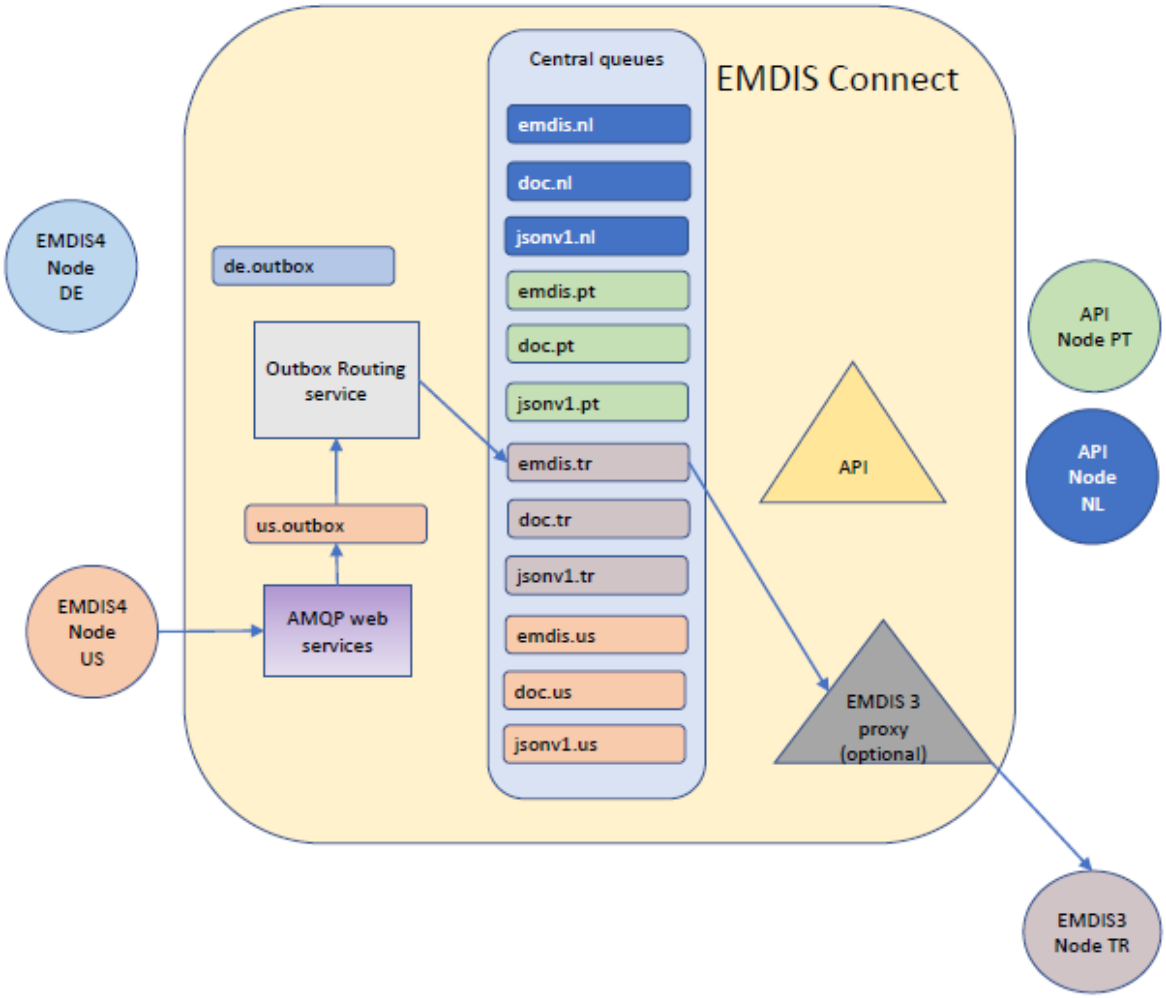
- emdis.x queues contain:
  - msg
  - meta
- doc.x queues contain:
  - "EMDIS4" document messages in JSON format
- Jsonv1.x queues contain messages stored in new format for:
  - Internal use by APIs
  - Later use by future AMQP users
- Requires message type and version to be part of meta info of a message
- Most likely clients cannot specifically retrieve doc messages from queue à needs separate queue from other jsonv1 in case a message comes in from API and is awaiting translation

From: EMDIS 4 (US)
To: EMDIS 3 (TR)
WITH "inbox" & "outbox"



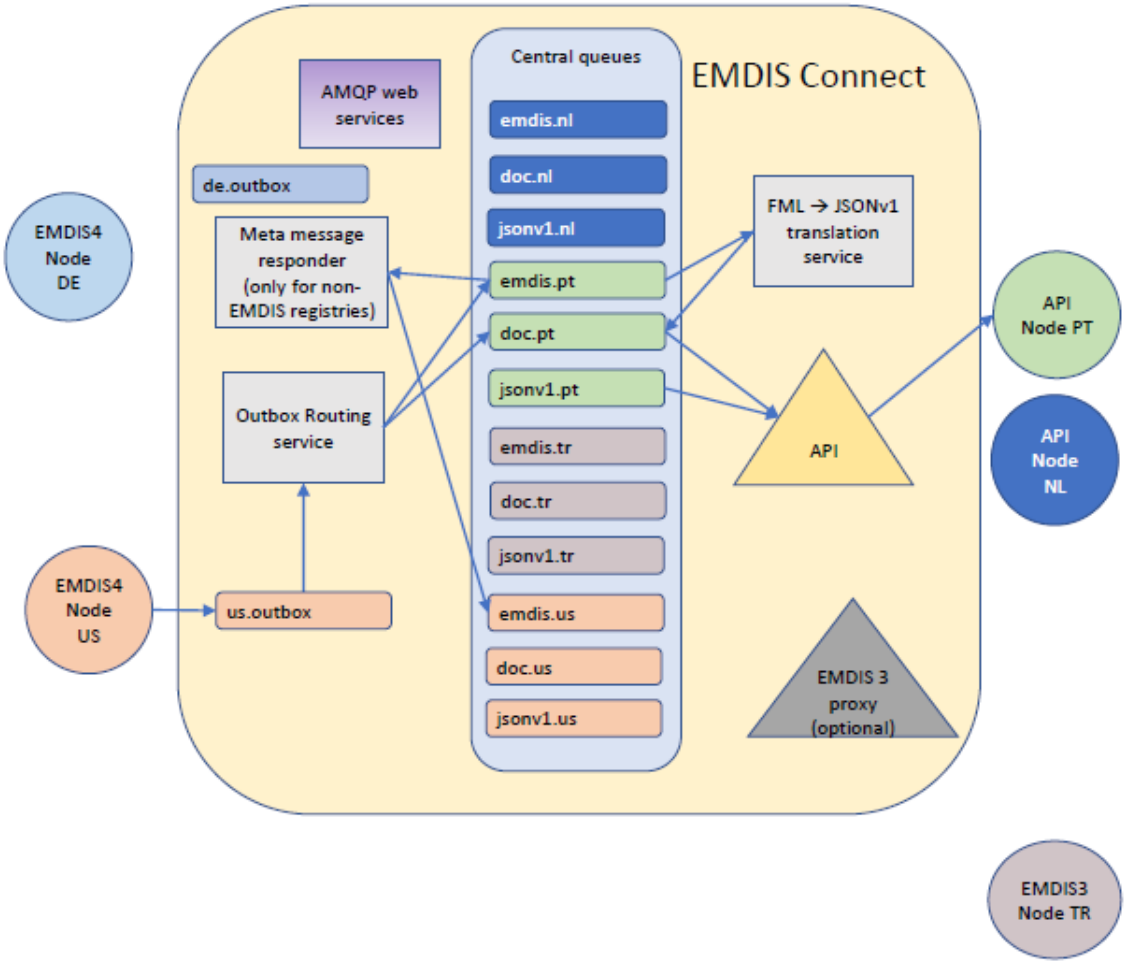
- emdis.x queues contain:
  - msg
  - meta
- doc.x queues contain:
  - "EMDIS4" document messages in JSON format
- Jsonv1.x queues contain messages stored in new format for:
  - Internal use by APIs
  - Later use by future AMQP users
- Requires message type and version to be part of meta info of a message
- Most likely clients cannot specifically retrieve doc messages from queue à needs separate queue from other jsonv1 in case a message comes in from API and is awaiting translation

From: EMDIS 4 (US) web sockets
To: EMDIS 3 (TR)
WITH "inbox" & "outbox"



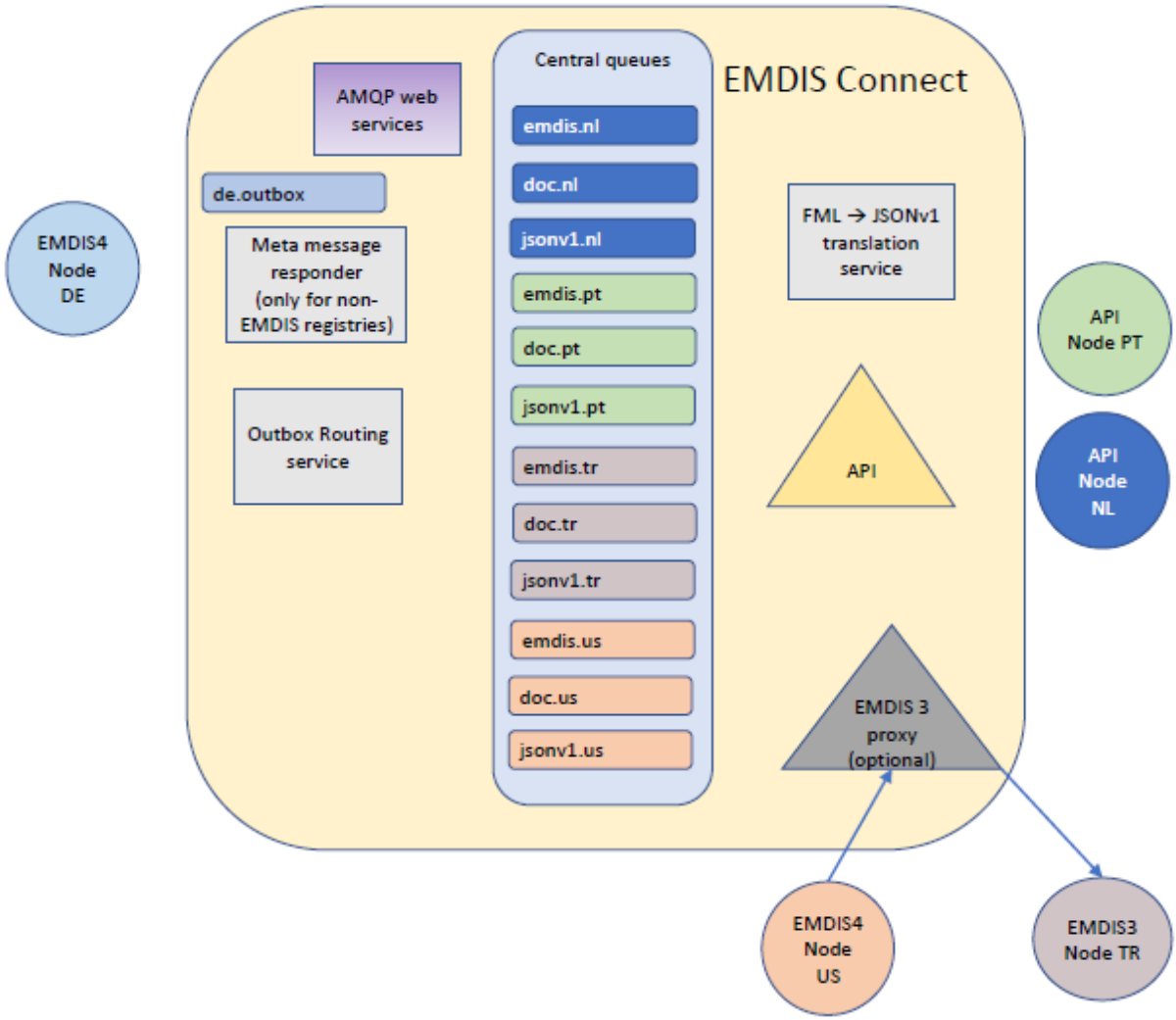
- emdis.x queues contain:
  - msg
  - meta
- doc.x queues contain:
  - "EMDIS4" document messages in JSON format
- Jsonv1.x queues contain messages stored in new format for:
  - Internal use by APIs
  - Later use by future AMQP users
- Requires message type and version to be part of meta info of a message
- Most likely clients cannot specifically retrieve doc messages from queue à needs separate queue from other jsonv1 in case a message comes in from API and is awaiting translation

From: EMDIS 4 (US)
To: API (PT)
WITH "inbox" & "outbox"



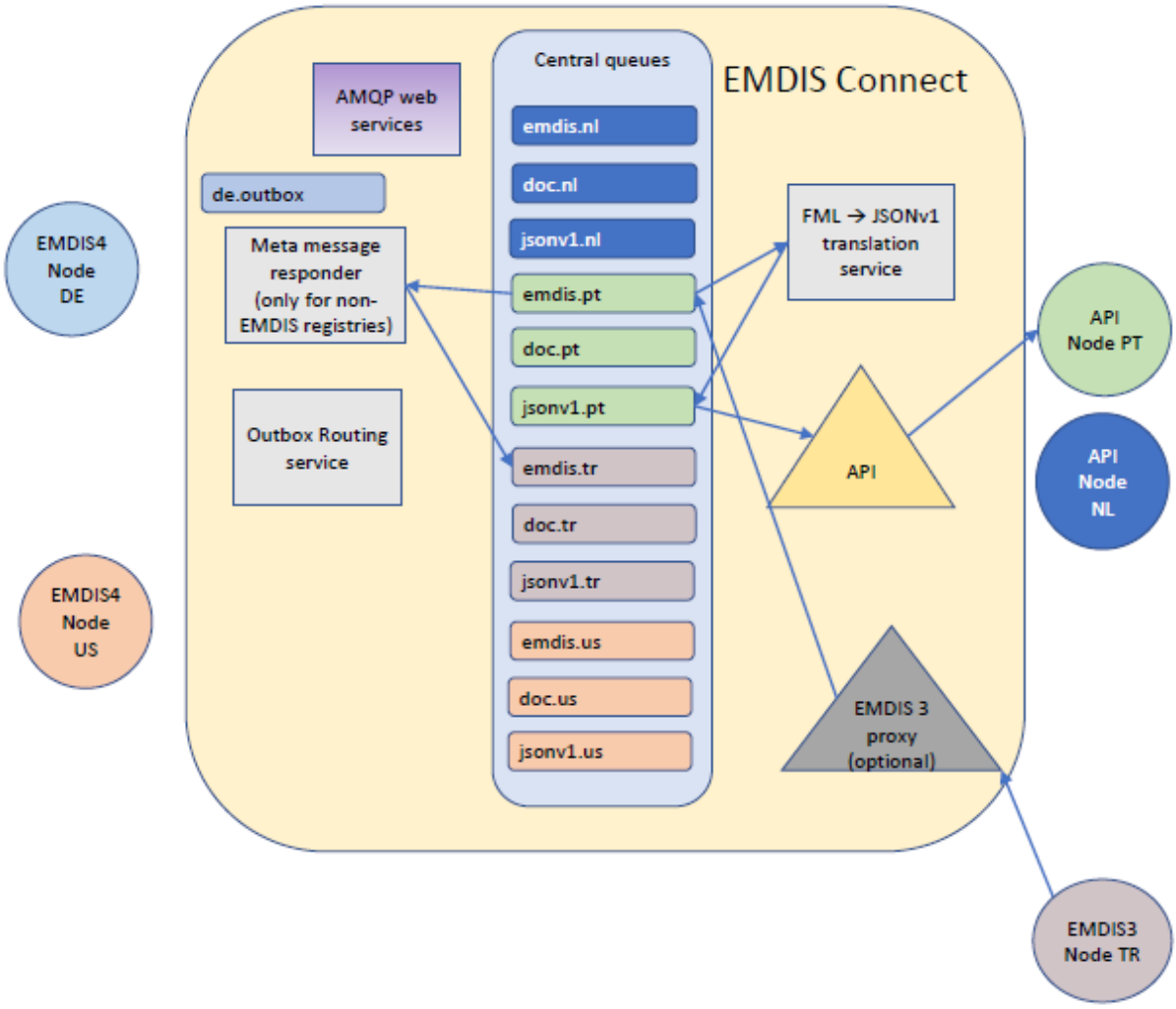
- emdis.x queues contain:
  - msg
  - meta
- doc.x queues contain:
  - "EMDIS4" document messages in JSON format
- Jsonv1.x queues contain messages stored in new format for:
  - Internal use by APIs
  - Later use by future AMQP users
- Requires message type and version to be part of meta info of a message
- Most likely clients cannot specifically retrieve doc messages from queue à needs separate queue from other jsonv1 in case a message comes in from API and is awaiting translation

From: EMDIS 3 (US)
To: EMDIS 3 (TR)
WITH "inbox" & "outbox"



- emdis.x queues contain:
  - msg
  - meta
- doc.x queues contain:
  - "EMDIS4" document messages in JSON format
- Jsonv1.x queues contain messages stored in new format for:
  - Internal use by APIs
  - Later use by future AMQP users
- Requires message type and version to be part of meta info of a message
- Most likely clients cannot specifically retrieve doc messages from queue à needs separate queue from other jsonv1 in case a message comes in from API and is awaiting translation

From: EMDIS 3 (US)
To: EMDIS 3 (TR)
WITH "inbox" & "outbox"



- emdis.x queues contain:
  - msg
  - meta
- doc.x queues contain:
  - "EMDIS4" document messages in JSON format
- Jsonv1.x queues contain messages stored in new format for:
  - Internal use by APIs
  - Later use by future AMQP users
- Requires message type and version to be part of meta info of a message
- Most likely clients cannot specifically retrieve doc messages from queue à needs separate queue from other jsonv1 in case a message comes in from API and is awaiting translation

Appendix 2 – Educational webinar on Alternative Algorithms

matching donors • serving patients

WMDA's Search, Match & Connect:
Alternative Algorithms

presented by Alicia Venter, Mark Melchers, Jan Hofmann,
Christine Urban and Zabeen Patel

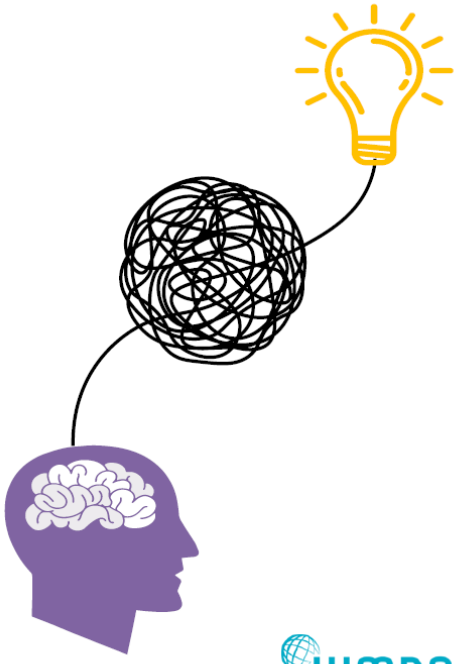
Outline

1. Background
2. What is an algorithm?
3. Current algorithm in use
4. Haplotype Frequency sets
5. Why do we need an alternative?
6. HAP-E Search & ATLAS
7. Comparative examples
8. Take home messages

What is an algorithm and how does it work?

In essence, **algorithms** are simply a series of instructions that are followed, step by step, to **do** something useful or **solve a problem**.

For example, you could consider a cake recipe an **algorithm** for making a cake. In computing, **algorithms** provide computers with a successive **guide to completing actions**.



6

Current algorithm in use - **Optimatch**

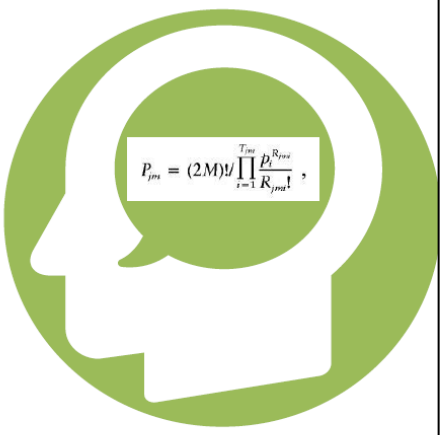
The OptiMatch matching algorithm helps to provide probability matching using haplotype frequencies.

WMDA collects the data from donors and cords on behalf of the listing organisations and OptiMatch is connected to this database.

The data collected is used to periodically calculate Haplotype Frequency (HF) sets which are used by the algorithm to perform the probability matching.

Haplotype Frequency (HF) sets

- HF is defined as the proportion of chromosomes of that type in the population.
- WMDA uses HF set configuration to better serve all patients and provide increased accuracy when matching donors and cord blood units to patients.
- These frequency sets utilize high resolution typing results from organisations to extrapolate the haplotypes of the region or organisation.
- Thus, an organisation or geographical region must meet a minimum threshold of number of donors and availability of high resolution typing to build usable and valuable frequency set.
- For organisations or geographical regions that do not meet this minimum threshold, the global HF set is applied.
- To check which HF set is applied to your organisations donors and/or CBUs, please visit this link
- It is important to know that your organisations frequency set is changeable! You can request that your donors be returned to the global consensus HF set or to a regional set.



9

IMPORTANT!!!

Probability matching will just calculate chances based on assigned HF sets and you should use it as a tool in combination with your own knowledge of HLA.

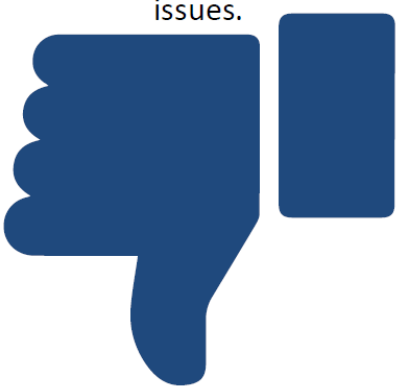
Do not select donors/cords solely on the probabilities calculated by the system.

10

Why do we need Alternative Algorithms?

The Search & Match User Interface, API and future search tools will stay the same as much as possible to help ease the switch between algorithms.

Cannot offer Optimatch via the Search API and other future search tools due to licensing issues.



Why do we need Alternative Algorithms?

Hap-E Search (DKMS) and ATLAS (Anthony Nolan) is currently in development

CHOICE

Some algorithm might work better for certain situations, such as mismatch searches.

CONTINUITY

If there are issues with a certain matching algorithm, WMDA can offer alternatives, ensuring that you will always be able to find the right donor/CBU for your patients.

CURRENT

With future development of "side-loading", donor information about non-HLA criteria is loaded in real-time and is not affected by the matching algorithm.

HAP-E & ATLAS

FEATURES

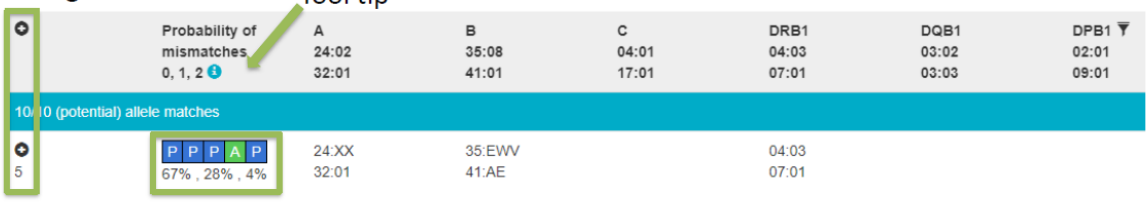
- Up to **five loci**, allelic-level HLA matching
- DPB1 "**permissive mismatch**" scoring
- **Match prediction** across the haplotype as well as per locus
- **No restriction on number of matches** returned from algorithm
  - Only first 2000 visible in S&M

Optimatch vs HAP-E vs ATLAS

SIMILARITIES

Ranking

Tool tip

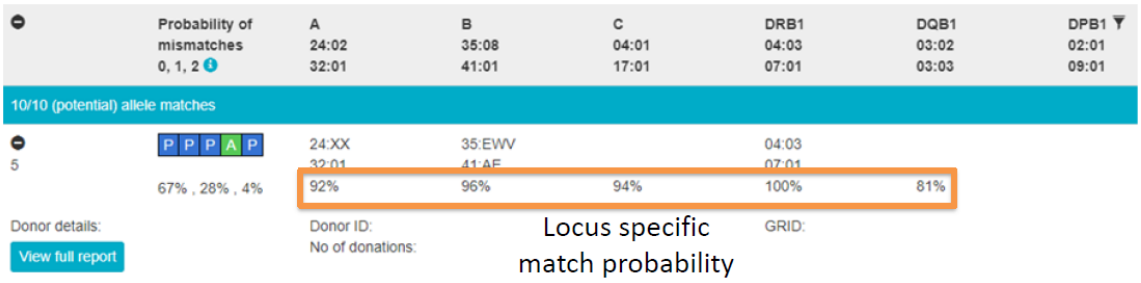


Letter/Colour match code

Probability of 0, 1, 2 mismatches

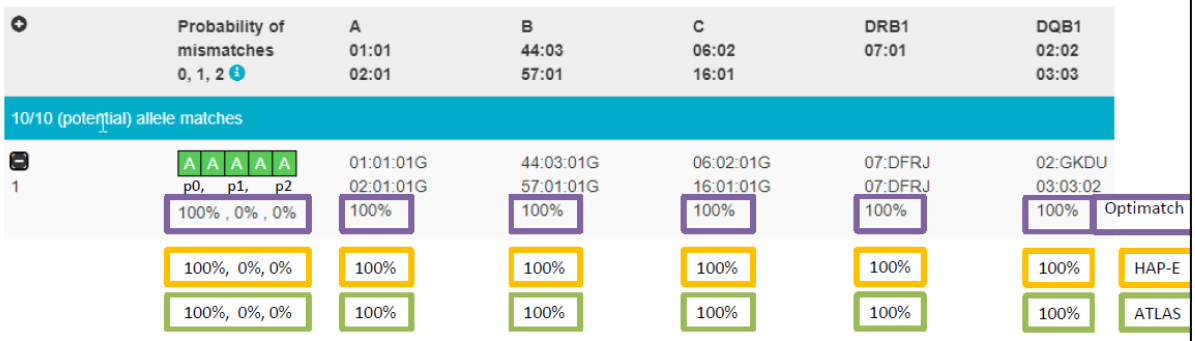
- Regardless of the algorithm applied, the ranking remains as described in S&M User Guide
- The letter/colour match code
- Probabilities of 0, 1 and 2 mismatches
- Tool tip regarding probability of mismatches calculation
- High probabilities predict matches and low probabilities predict mismatches

Optimatch vs HAP-E vs ATLAS
DIFFERENCES



- Optimatch and ATLAS calculate the probability of a mismatch based on all haplotypes and all numbers of mismatches
- For locus-specific matching probabilities, HAP-E calculates the probability of a match considering ONLY the known + 1 additional mismatch on the whole haplotype
- Differences between the calculations will only become apparent in cases of a mismatch

Example 1



Example 2

	Probability of mismatches 0, 1, 2	A 01:01P 01:01:01G	B 08:01:01G 44:03:02	C 07:02:01G 07:01:01G	DRB1 03:01:01G 03:01:01G	DQB1 02:01:01 02:01:01	
782	P M P P P p0, p1, p2	01:APUZU 01:APUZU	(07:JXWY) 08:JWZE		03:KBVN 03:KBVJ		
	0% , 76% , 24%	99%	0%	76%	99%	99%	Optimatch
	0% , 76%, 24%	100%	100%	76%	100%	100%	HAP-E
	0% , 76%, 24%	99%	99%	76%	99%	99%	ATLAS

The donor already has a known mismatch. HAP-E indicates the probability that the locus will be a match for up to one additional mismatch (so the combined 9/10 and 8/10 case).
➔For the known mismatch locus (B*), HAP-E does not consider the allele that is already known to be a mismatch

Alternative Matching Algorithms

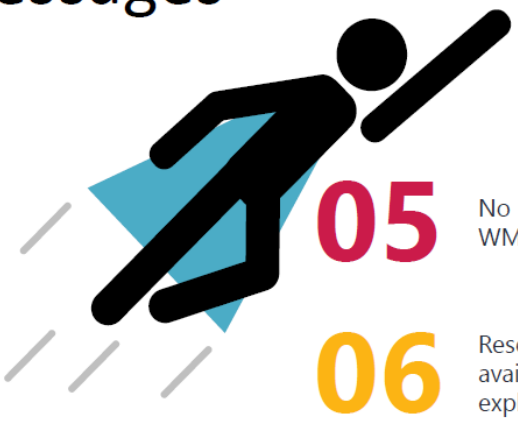
Validation
Against existing trusted algorithms

Compatible
With various internal systems and APIs

Cloud Efficient
Compatible with MS Azure

No Major Feature Loss
WMDA Search & Match Service should appear relatively unchanged

Take Home Messages



01 Integration in 2020
Testing and launch in 2021

02 No algorithm bias, calculations independently verified, ranking in S&M

03 There are more similarities than there are differences

04 HAP-E and ATLAS assist you in predicting the location of additional mismatches

05 No cost implication to WMDA members

06 Resources will be made available at launch to explain technical nuances

07 Freedom to choose which algorithm to apply. Optimatch will still be available until end 2021

08 Optimatch will NOT be applied to search results generated by APIs